JOURNAL OF RESEARCH IN SCIENCE TEACHING

Research Article

Investigating a Learning Progression for Energy Ideas From Upper Elementary Through High School

Cari F. Herrmann-Abell and George E. DeBoer

Project 2061, American Association for the Advancement of Science, 1200 New York Ave. NW, Washington, DC 20005

Received 9 September 2016; Accepted 7 June 2017

Abstract: This study tests a hypothesized learning progression for the concept of energy. It looks at 14 specific ideas under the categories of (i) Energy Forms and Transformations; (ii) Energy Transfer; (iii) Energy Dissipation and Degradation; and (iv) Energy Conservation. It then examines students' growth of understanding within each of these ideas at three levels of increasing conceptual complexity. The *basic* level of the model focuses on simple energy relationships and easily observable effects of energy processes; the *intermediate* level focuses on more complex energy concepts and applications; and the *advanced* level focuses on still more complex energy concepts, often requiring an atomic/molecular model to explain phenomena. The study includes results from 359 distractor-driven, multiple-choice test items administered to over 20,000 students in grades 4 through 12 from across the U.S. Rasch analysis provided linear measures of student performance and item difficulty on the same scale. Results largely supported a model of students' growth of understanding that progresses from an understanding of forms and transformations of energy to energy transfer to conservation while also progressing along a separate dimension of cognitive complexity. An analysis of the current state of students' understanding with respect to the knowledge identified in the learning progression showed that elementary level students perform well in comparison to expectations but that middle and high school students' performance does not meet expectations. J Res Sci Teach 9999:XX–XX, 2017

Keywords: energy; assessment; Rasch analysis; learning progressions

With the publication of *A Framework for K-12 Science Education* (National Research Council [NRC], 2012) and the *Next Generation Science Standards* [NGSS] (NGSS Lead States, 2013), the focus on learning progressions has taken a more prominent role in science education research. The NRC, in *A Framework for K-12 Science Education,* summarizes the role of learning progressions in science education as follows:

> To develop a thorough understanding of scientific explanations of the world, students need sustained opportunities to work with and develop the underlying ideas and to appreciate those ideas' interconnections over a period of years rather than weeks or months. This sense of development has been conceptualized in the idea of learning progressions. If mastery of a

Contract grant sponsor: Institute of Education Sciences; Contract grant number: R305A120138.
Correspondence to: C. F. Herrmann-Abell; E-mail: cabell@aaas.org
DOI 10.1002/tea.21411
Published online in Wiley Online Library (wileyonlinelibrary.com).

> core idea in a science discipline is the ultimate educational destination, then well-designed learning progressions provide a map of the routes that can be taken to reach that destination (NRC, 2012, p. 26).

Such road maps are based on an examination of the structure of knowledge in a particular domain as well as on research on how students learn in that domain. Inevitably, any learning progression that is described must be a distilled version of the incredibly complex network of associated ideas and paths that individual students take as they move toward an understanding of science ideas. But, even so, these learning progressions have the potential to better organize instruction, curriculum, and assessment across grade bands by moving away from conceptualizing science as discrete pieces of knowledge and toward a more coherent structure organized around a focused set of core ideas (NRC, 2007). The learning progression approach brings attention to where the students are coming from and where they currently are in their development of science understanding in order to better help them move along the progression on the way to science literacy. When paired with formative assessments, learning progressions become powerful tools for teachers to use to diagnose gaps in understanding and to inform the development of trajectories for future instruction (Heritage, 2008).

Researchers have described learning progressions for physical science topics such as matter (e.g., Hadenfeldt, Neumann, Bernholt, Liu & Parchmann, 2016) and energy (e.g., Neumann, Viering, Boone, & Fischer, 2013), earth science topics such as the water cycle (Forbes, Zangori, & Schwarz, 2015) and climate change (e.g., Breslyn, McGinnis, McDonald, & Hestness, 2016), and life science topics such as genetics (e.g., Todd & Kenyon, 2015) and ecosystems (e.g., Hokayem & Gotwals, 2016). In addition to learning progressions in these content areas, researchers have also described learning progressions for science practices such as argumentation (e.g., Osborne et al., 2016).

For our purpose, learning progressions are descriptions of the order in which ideas that comprise a content domain are most likely to be effectively learned. They describe a continuum of successively more sophisticated ways of thinking about a concept that develops over time (American Association for the Advancement of Science [AAAS], 2001, 2007; Corcoran, Mosher, & Rogat, 2009; NRC, 2007). The upper level, or "upper anchor," of a learning progression specifies the knowledge that instruction is ultimately building toward and that students are expected to have in order for them to be considered proficient in that area. The lower levels identify productive "steps along the way" that students should follow on the path to proficiency.

The starting point of a learning progression is typically referred to as its "lower anchor." For their elementary grades learning progression (grades 3, 4, and 5), Lacy, Tobin, Wiser, and Crissman (2014) used the knowledge that students come to third grade with as their lower anchor. Although it is true that good instruction must take into account all the ideas that young children bring to school, including their misconceptions, for assessment purposes, we focused our lower anchor on the correct ideas that students are expected to have by the end of elementary school.

When creating a learning progression, one begins by considering the logical structure of the relevant disciplinary domain (i.e., the fact that some ideas necessarily depend on others) as well as the available research on students' learning. Once articulated, the hypothesized progression is empirically validated. Typical validation approaches include either (i) classroom interventions to determine what students are capable of learning or (ii) cross-sectional studies that portray the current status of what students at different levels know (Duncan & Hmelo-Silver, 2009).

The current study falls into the second category. We are not testing whether good instruction from elementary through high school will produce the desired results. Instead, we are testing the progression of difficulty of these ideas as indicated by student performance in an environment of

typical instruction. We use the increasing difficulty of the ideas as an indicator of the order in which the ideas are learned. The study is based on student scores on 359 multiple choice assessment items across the energy ideas in our proposed learning progression.

Energy is a central topic in the K-12 science curriculum, with many applications in the earth, physical, life sciences, and in engineering and technology. Therefore, it is important to know how students' thinking about energy develops so that they can be appropriately supported in their understanding of energy. This study tests the order in which four broad categories of energy ideas, generally considered to comprise the energy concept (Duit, 2014), are learned: (i) Energy Forms and Transformations; (ii) Energy Transfer; (iii) Energy Dissipation and Degradation; and (iv) Energy Conservation. It then examines students' growth of understanding of more specific ideas within each of these categories and across three levels of increasing conceptual complexity.

A number of studies have investigated learning progressions for the energy concept (Liu & Collard, 2005; Lee & Liu, 2010; Liu & McKeough, 2005). Liu and McKeough (2005) used the responses from three populations of U.S. students (3rd and 4th graders, 7th and 8th graders, and 12th graders) to 27 multiple-choice and short-answer items from the TIMSS database. In a follow up study, Liu and Collard (2005) administered three performance assessments to 67 students from one 4th grade class, one 8th grade class, and one high school physics class in the U.S. Lee and Liu (2010) selected eight multiple-choice items and two explanation items from item sets released by TIMSS and the National Assessment of Educational Progress (NAEP) and tested them with 2,688 middle school students from across the U.S. Each of these studies concluded that students' understanding of energy progressed through four conceptual categories. First, students perceive energy as activity or the ability to do work. As students' understanding grows, they begin to distinguish different energy sources and forms of energy. Next comes an understanding of energy transfer, followed by an awareness of energy degradation. Finally, at the upper level of the progression, students are able to accept the highly abstract idea of conservation of energy.

The approach that has been taken by researchers to validate this energy progression is to compare the relative difficulty of the four energy categories. More recently, researchers have been investigating students' growth of understanding within each category as a way to fine tune the progression. This is typically done by looking at conceptual complexity as a separate dimension on which progress can be observed within the content categories. For example, Neumann et al. (2013) designed an assessment that tested a progression of complexity within each of four energy categories (i.e., forms, transfer, degradation, etc.), starting with students' understanding of facts, then moving to simple connections, to qualified relationships, and finally to complex concepts. They administered this assessment to 1,856 German students in 6th through 10th grades. Although their results did not support their proposed progression of conceptual complexity within each energy category, the results did show that students' understanding progressed in a series of overlapping rather than discrete steps through the four energy categories This suggests that students make progress by understanding aspects of multiple and interrelated energy concepts at the same time, not by mastering one concept before moving on to the next.

The idea that students make progress on multiple interconnected pathways and not in a simple linear way is not surprising. Although learning progressions may seem to imply a linear sequence, with each subsequent idea in the progression building on each previous idea, knowledge is much more complex than that and is better characterized as multiple interwoven strands that create complex networks of ideas. And student learning of these ideas adds another layer of complexity because of the differences in the experiences that each student brings to the classroom and how students create knowledge from those varied experiences. A number of approaches to ensuring that students learn in this interconnected way include an emphasis on curriculum coherence (Roseman, Stern, & Koppal, 2010) and knowledge integration (Linn, 2006).

Any proposed learning progression should acknowledge this complexity, both in how the upper anchor is envisioned and how the building blocks or stepping stones reflect the network of interconnected ideas that lead to that upper anchor. For example, the idea that there are different ways in which energy manifests itself is helpful in understanding that energy can be transferred from one place to another, as when a wood fire is used to heat the air in a room. The thermal energy of the air had to come from somewhere. It came from a chemical reaction between the wood and oxygen in which chemical energy was transformed into thermal energy and transferred to the air. As a second example, conservation of energy may seem counterintuitive without understanding that energy can be transferred or transformed and that the dissipation of energy to the surrounding environment accompanies all energy transfers and transformations. In fact, the idea of conservation of energy has been specifically identified as one that requires a high level of knowledge integration (Goldring & Osborne, 1994; Lacy et al., 2014; Lee & Liu, 2010).

How Next Generation Science Standards (NGSS) Treats the Energy Concept

When laying out the ideas that students should learn about energy and the sequence in which they should be learned, the *Framework for K-12 Science Education* and NGSS used a similar conceptual structure to that described by other researchers referenced in this paper. The *Framework*, in particular, is clear that the end goal is that students appreciate that a system's total energy is conserved unless energy enters or leaves the system. When it appears that energy has been lost, it is because energy has left the system even if the amount is small. And, although not explicitly stated, the *Framework* and NGSS generally begin with concrete and familiar contexts for elementary school students and move to more abstract and less familiar contexts in high school. By examining the sequence of ideas in NGSS and the *Framework*, it is also clear that the writers of those documents believe that students should learn aspects of the energy concept in an integrated manner throughout the grade bands, beginning in elementary school.

Compared to the work of other researchers, however, there are a number of places where the NGSS story is not as complete as it could be. The most notable example is that NGSS does not include the idea of dissipation at the elementary and middle grades (even though it is included at both those levels in the *Framework*). In other places an idea may be presented in elementary school and then not carried out through middle and high school, or an idea appears only at the high school level without having been introduced earlier. A full comparison of the way that NGSS treats the energy concept compared to what is proposed in this paper can be seen in Table 2. The learning progression that was tested in the work reported on here begins with the four major categories of energy concepts that other researchers have described (forms and transformations; transfer; dissipation and degradation; and conservation), and then elaborates on this conceptual structure by including five specific energy forms and six specific modes of energy transfer. Finally, it formalizes the use of concrete and familiar contexts at the early grades and abstract, often atomic/molecular contexts, at the upper level. The goal was to create and then test a fuller description of the energy construct than had been previously described, and to systematically vary the conceptual complexity within each idea. This was all possible because of the large number of items that we had developed (359) and the large number of students we were able to test (over 20,000).

Our research had two main purposes. The first was to test the validity of the comprehensive progression of understanding of energy described in this paper. The second was to determine the current state of students' understanding of that energy concept at three grade levels—upper elementary, middle, and high school. The study sought to answer the following specific questions:

1) To what extent do the results of our study support the currently established learning progression for energy across four broad categories of energy concepts?
2) To what extent do the results of our study support the currently established learning progression across four broad energy categories when data are analyzed at the level of specific ideas within those categories?
3) To what extent do the results of our study support a hypothesized progression of understanding across three levels of conceptual complexity for each of the specific energy ideas?
4) How are students currently performing at each grade band with respect to the expectations described in the hypothesized learning progression?

Methodology

Defining the Construct for an Energy Learning Progression

As already noted, the concept of energy is typically separated into four categories: (i) Energy Forms and Transformations, the idea that energy manifests itself in different forms, such as kinetic energy and gravitational potential energy, that can be converted from one to another; (ii) Energy Transfer, the idea that energy can be transferred from one location to another in different ways; (iii) Energy Dissipation and Degradation, the idea that whenever energy is transformed or transferred some energy is also transferred to the environment as thermal energy; and (iv) Energy Conservation, the idea that the total amount of energy in a system remains constant unless energy is added to or released from the system. It was on those four broad conceptual categories that student understanding was assessed.

For two of the categories—Energy Forms and Transformations and Energy Transfer—we further defined the specific ideas that make up those categories. For the Energy Forms and Transformations category, we identified and assessed student understanding of five forms of energy along with the idea of energy transformation itself, and we expanded the Energy Transfer category into six specific mechanisms of energy transfer. The forms of energy include (i) kinetic energy, the energy associated with motion; (ii) thermal energy, the energy associated with temperature; (iii) gravitational potential energy, the energy associated with distance from the center of the earth; (iv) elastic potential energy, the energy associated with the stretching, bending, or twisting of an elastic object; and (v) chemical energy, the energy associated with arrangements of atoms in a chemical reaction system. Energy Transformations, that is, the conversion of one of these forms of energy into another, makes up the sixth idea in this category. The Energy Transfer category includes (i) conduction, the transfer of energy due to temperature differences between objects in contact; (ii) convection, the transfer of energy due to the movement of liquids or gases; (iii) radiation, the transfer of energy by electromagnetic waves; (iv) mechanical energy transfer, the transfer of energy by forces exerted by one object on another; (v) the transfer of energy by sound; and (vi) electrical transfer, the transfer of energy in a complete electrical circuit. This gives us a total of 14 specific ideas in our energy construct: five forms of energy ideas, one energy transformation idea, six energy transfer ideas, one energy dissipation/degradation idea, and one conservation of energy idea (see Table 1).

For each of the energy ideas described above, three levels of conceptual complexity were specified. At the *basic* level, students were expected to be able to think about the most easily observable aspects of energy—objects with more thermal energy are warmer, objects with more motion energy move faster—and to recognize obvious effects of simple energy processes—a rock dropped from a greater height will do more damage than one dropped from a lower height. At the

Table 1
Energy ideas targeted by the assessment items

Ideas About the Forms of Energy	Ideas About Energy Transfer	Other Energy Ideas
Kinetic energy	Transferring energy by conduction	Energy conservation
Thermal energy	Transferring energy by convection	Energy dissipation & degradation
Gravitational potential energy	Transferring energy by radiation	
Elastic potential energy	Transferring energy by forces	
Chemical energy	Transferring energy electrically	
Energy transformations	Transferring energy by sound	

next level, the *intermediate* level, students were expected to be familiar with less easily observable aspects of energy—thermal energy is related to both temperature and mass—and to be able to explain energy-related phenomena or evaluate energy applications using more complex energy concepts. At the highest level, the *advanced* level, students were expected to understand even more complex and abstract energy concepts, often requiring an atomic/molecular model to explain phenomena. For example, students were expected to know that the thermal energy of an object also depends on the random motion of its atoms and molecules.

Many energy ideas can easily be placed into three distinct levels of conceptual complexity. For example, at the basic level students can be expected to know that the motion energy of an object is related to its observable speed; at the intermediate level they can be expected to know that the motion energy of an object is related to its mass as well as its speed; and at the advanced level, they can be expected to know that the relationship between speed, mass, and motion energy is non-linear. In the case of conduction, at the basic level students can be expected to know that when a warmer object is placed in contact with a cooler object, the warmer object will get cooler and the cooler object will get warmer. At the next level they can be expected to know that conduction occurs because energy is transferred from the warmer object to the cooler one. At the highest level students can be expected to know that this energy is transferred by the random collisions of atoms and molecules that make up the objects. For gravitational potential energy, at the basic level idea students can be expected to know that the higher an object is above the earth, the more energy it has and the more impact it will have when dropped. At the next level students can be expected to know that for objects near the surface of the earth, gravitational potential energy depends on the distance the object is above the earth and the mass of the object. At the highest level students can be expected to know that gravitational potential energy is associated with the separation of mutually attracting masses.

In summary, our hypothesized energy learning progression predicts growth in student understanding along a continuum of conceptual complexity that moves from: (i) an awareness of easily observable energy phenomena and the application of basic energy ideas to explain events in the world; to (ii) the use of more complex energy concepts to explain phenomena; to (iii) the use of advanced energy concepts to explain less easily observable phenomena, often requiring an atomic/molecular explanation. Descriptions of the progressions of understanding for each idea tested in this study are presented in Table 2. Note that for the Transferring Energy Electrically idea, we created only two levels in the progression, and for the Energy Transformations idea, only one level. The knowledge statements in Table 2 were drawn from *Benchmarks for Science Literacy* (American Association for the Advancement of Science [AAAS], 1993), *Atlas of Science Literacy* (AAAS, 2001, 2007), *A Framework for K-12 Science Education* (NRC, 2012), and *Next Generation Science Standards* (NGSS Lead States, 2013).

Table 2
Proposed progression of understanding for energy ideas and alignment with NGSS Performance Expectations (PEs)

Energy Idea	Basic Level	Intermediate Level	Advanced Level
Kinetic energy	The amount of energy an object has depends on how fast it is moving.	The kinetic energy (motion energy) of an object depends on the speed and the mass of the object.	Kinetic energy (motion energy) is proportional to the mass of a moving object and increases rapidly with increasing speed.
PEs	4-PS3-1	MS-PS3-1[a]	MS-PS3-1[a]
Thermal energy	The amount of energy an object has depends on how warm it is.	The thermal energy of an object depends on the temperature and the mass of the object and the material of which the object is made.	The thermal energy of an object depends on the disordered motions of its atoms or molecules and the number and types of atoms or molecules of which the object is made.
PEs	–	MS-PS3-4	MS-PS1-4, HS-PS3-2
Gravitational potential energy	The amount of energy an object has depends on how high it is above the surface of the earth.	The gravitational potential energy of an object near the surface of the earth depends on the distance the object is above the surface of the earth (or an alternate reference point), and the mass of the object.	Gravitational potential energy is associated with the separation of mutually attracting masses.
PEs	–	MS-PS3-2	HS-PS3-2
Elastic potential energy	The amount of energy an elastic object has depends on how much the object is stretched, compressed, twisted, or bent.	The elastic potential energy of an elastic object depends on the amount the object is stretched or compressed and how difficult it is to stretch or compress the object.	The amount of elastic potential energy stored in an elastic object increases when the object is stretched or compressed because stretching and compressing an object changes the distances between the atoms and molecules that make up the object.
PEs	–	–	HS-PS3-2
Chemical energy	Energy is released when fuel is burned. Energy is also released when food is used as fuel in animals.	Some chemical reactions release energy into the surroundings, whereas other chemical reactions take in energy from the surroundings.	Chemical energy is associated with the arrangement of atoms that make up the molecules of the reactants and products of a chemical reaction. Because the arrangement of atoms making up the molecules is different before and after the chemical reaction takes place, the amount of chemical energy in the system is also different.
PEs	5-PS3-1	–	HS-PS1-4

continued

Energy Idea	Basic Level	Intermediate Level	Advanced Level
Energy transformations		Most of what goes on in the universe—from exploding stars and biological growth to the operation of machines and the motion of people—involves some form of energy being converted into one or more other forms of energy.	
PEs		4-PS3-4	HS-PS3-3
Transferring energy by conduction	When warmer things are touching cooler ones, the warmer things get cooler and the cooler things get warmer until they all are the same temperature.	Conduction is the transfer of energy that occurs when a warmer object (or sample of matter) comes in contact with a cooler object (or sample of matter) without a transfer of matter.	Energy is transferred by conduction through a material by the random collisions of atoms and molecules that make up the material.
PEs	4-PS3-2	MS-PS3-3	MS-PS3-4
Transferring energy by convection	When air or water moves to another location, it can change the temperature at that location.	Temperature variations in fluids such as air and water lead to currents that circulate the fluid and transfers energy from place to place in the fluid.	In a fluid, regions that have different temperatures have different densities. The differences in density lead to an imbalance between the downward gravitational force and upward (buoyant) forces exerted by the surrounding fluid, creating currents that contribute to the transfer of energy.
PEs	–	–	MS-ESS2-6
Transferring energy by radiation	When light shines on an object, the object typically gets warmer.	Light transfers energy from a light source to a receiver.	Energy can be transferred by electromagnetic radiation.
PEs	K-PS3-1, K-PS3-2	4-PS3-2	–
Transferring energy by forces	Pushes and pulls can transfer energy from one object to another resulting in a change in the objects' motion.	Energy is transferred mechanically whenever an object exerts a force, either by contact or at a distance, on another object that changes the objects' position or shape.	When two objects change relative position as a result of a gravitational, magnetic, or electric force, the potential and kinetic energies of the system change.
PEs	4-PS3-3	MS-PS3-2	HS-PS3-5
Transferring energy electrically	Energy can be transferred electrically when an electrical source is connected in a complete circuit to an electrical device.		Electrostatic potential energy can be stored in the separation of charged objects.
PEs	4-PS3-2		HS-PS3-5
Transferring energy by sound	Sound can transfer energy from one location to another.	Energy can be transferred by sound when a vibrating object produces sound that travels through a medium to a receiver.	Energy is transferred by sound because of coordinated collisions between the atoms or molecules that make up the medium through which the sound travels.

continued

Energy Idea	Basic Level	Intermediate Level	Advanced Level
PEs	4-PS3-2	–	–
Energy dissipation & degradation	Objects tend to get warmer when they are involved in energy transfers.	Transformations and transfers of energy within a system usually result in some energy being released into its surrounding environment causing an increase in the thermal energy of the environment.	Unless prevented from doing so, energy will become uniformly distributed.
PEs	–	–	HS-PS3-4
Conservation of energy	Everything has energy.	A decrease in energy in one object or set of objects always is accompanied by an increase in energy in another object or set of objects.	Regardless of what happens within a system, the total amount of energy in the system remains the same unless energy is added to or released from the system.
PEs	–	MS-PS3-5	HS-PS3-1

[–]Indicate that no PEs match that particular level of the progression.
[a]MS-PS3-1 aligns to the intermediate kinetic energy level while the underlying disciplinary core idea aligns to the advanced kinetic energy level.

As we noted earlier, our learning progression begins with the four major categories typically used to describe the energy concept. The expectation is that students will use these ideas in progressively more sophisticated ways to develop a coherent, integrated understanding of energy, its unitary nature, and its conservation. We also noted that this is the basis for the specification of learning goals in the NRC's *A Framework for K-12 Science Education* and in NGSS. To be explicit about how closely the learning progression that we tested matches what is in the *Framework* and NGSS, we examined the NGSS performance expectations (PEs) listed for each grade and the underlying disciplinary core ideas (DCIs) found in the foundation boxes under each PE. For the NGSS energy core idea, there are seven elementary PEs, five middle school PEs, and five high school PEs. Additionally, we identified two middle school PEs (MS-PS1-4 and MS-ESS2-6) and one high school PE (HS-PS1-4) that were listed under other NGSS core ideas. We then matched each PE to our learning progression (see Table 2).

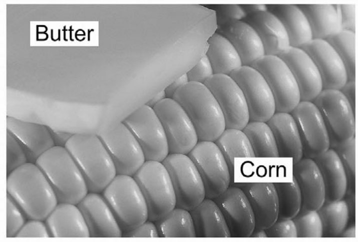
Overall we found very good alignment between the NGSS PEs and our learning progression. All of the 14 energy ideas could be matched with at least one PE. When we looked at the alignment by cognitive complexity level, we saw that, for the most part, our basic level of conceptual complexity matches the NGSS elementary school PEs; the intermediate level corresponds with the middle school PEs; and the advanced level parallels the high school expectations. There are some ideas for which we were unable to find PEs that match every level in the progression. For example, although there is an elementary school PE for the idea that energy is transferred by sound, there are no middle or high school PEs for this idea. And although the elastic potential energy, convection, and dissipation ideas are included at the high school level, they are not included at the elementary or middle school levels. Additionally, there is very little in NGSS about thermal energy, gravitational potential energy, or conservation of energy in the elementary school grade band. In our progression, we included statements at all levels for each energy idea in order to present a more complete picture of the nature of energy. Introducing each idea at a basic level supports younger students' progress toward the complex understanding expected in the high school performance expectations. In their energy progression for elementary students, Lacy and colleagues (2014) included basic ideas about gravitational potential energy, thermal energy,

Table 3
Item count by level of progression for each idea

Energy Category	Energy Ideas	Number of Items		
		Basic	Intermediate	Advanced
Forms of energy	Kinetic energy	5	27	8
	Thermal energy	3	19	18
	Gravitational potential energy	6	23	6
	Elastic potential energy	4	11	3
	Chemical energy	4	16	8
	Energy transformations		29	
Energy transfer	Transferring energy by conduction	4	18	4
	Transferring energy by convection	3	7	7
	Transferring energy by radiation	3	10	13
	Transferring energy by forces	4	13	6
	Transferring energy electrically	2		9
	Transferring energy by sound	2	3	7
Conservation of energy		5	5	23
Energy dissipation & degradation		6	10	5

Basic level conduction item

A cook heats up some corn. Then she cuts a piece of cold butter and places it on top of the hot corn.



What will happen to the temperature of the corn and the butter as soon as she puts the butter on top of the corn?

A. Both the corn and the butter will get cooler.
B. Both the corn and the butter will get warmer.
C. The corn will get cooler, and the butter will get warmer.
D. The corn will stay the same temperature, but the butter will get warmer.

Intermediate level conduction item

The temperature of a plastic block is 60°F, and the temperature of a metal block is 40°F. A student puts the plastic block on top of the metal block. Will the blocks ever reach the same temperature? Why or why not?

A. Yes, but only for a little while because the metal block will continue to get warmer and the plastic block will continue to cool
B. Yes, because energy will be transferred from the plastic block to the metal block until they reach the same temperature
C. No, because the temperature difference is not large enough for energy to be transferred
D. No, because the blocks are made of different materials

Advanced level conduction item

A person pours a hot drink into a cup and then places a room temperature spoon in the cup. After a while, the person notices that the handle of the spoon has gotten hotter. What caused the handle to get hotter?

A. Heat molecules from the hot drink are absorbed by the spoon. These heat molecules travel to the handle of the spoon, making the handle hotter.
B. The molecules that make up the hot drink are rubbing against each other harder than the molecules that make up the spoon. The rubbing creates new energy that flows through the spoon to the handle, making the handle hotter.
C. The hot drink causes the molecules of the spoon to speed up. These faster moving molecules then move to the handle of the spoon, causing the handle to get hotter.
D. The hot drink causes the molecules of the spoon to speed up. When these faster moving molecules collide with slower moving molecules, energy is transferred to the slower moving molecules. These collisions continue to occur throughout the spoon until they reach the handle, making the handle hotter.

Figure 1. Sample assessment items aligned to the learning progression for transferring energy by conduction.

dissipation, and conservation, and demonstrated that they can be attained by elementary students when appropriate instructional supports are in place.

Item Development

A total of 359 distractor-driven, multiple-choice assessment items (Sadler, 1998) were developed for use in this study to test students' understanding of energy. Table 3 presents the number of items aligned to each energy idea by level of conceptual complexity. Each item was designed to be aligned to a single conceptual complexity level and a single energy idea as described by the energy construct.

Item construction followed rigorous and iterative item development procedures that have been described in detail elsewhere (DeBoer, Herrmann-Abell, & Gogos, 2007; DeBoer et al., 2008; DeBoer, Lee, Husic, 2008; DeBoer, Herrmann-Abell, Wertheim, & Roseman, 2009; Herrmann-Abell & DeBoer, 2014). Briefly, the process includes (i) the identification of documented misconceptions (e.g., Driver, Squires, Rushworth, & Wood-Robinson, 1994), which are then used as distractors (Sadler, 1998); (ii) a careful evaluation of the items' alignment to the targeted ideas about energy and the targeted level of cognitive complexity; and (iii) a close examination of the items for their overall psychometric effectiveness. Rasch analysis (Rasch, 1980) was used throughout the item development process to monitor the items' psychometric properties (Bond & Fox, 2007; Liu & Boone, 2006; Boone, Staver, & Yale, 2014). Figure 1 shows sample items aligned to each level of the learning progression for the idea of Energy Transfer by Conduction. Additional sample items can be found in the supplementary materials and on our website, http://assessment.aaas.org/.

Alignment of items to the energy construct that we developed for the assessments was aided by the use of two criteria: *necessity* ensures that the targeted energy idea is needed to evaluate the answer choices, and *sufficiency* ensures that the targeted energy idea is enough by itself to answer correctly (Stern & Ahlgren, 2002). Careful alignment increases the validity of the inferences that can be made about what students know. To further ensure construct validity, the items and the energy construct were reviewed by a panel of scientists and science education experts to guarantee scientific accuracy of the construct and the items, and to eliminate construct irrelevant features such as issues with comprehensibility, test-wiseness, or inappropriate task contexts.

Items were pilot-tested with students in grades 4 through 12 to obtain feedback from them about whether the items were effectively measuring their understanding of the target learning goal (DeBoer et al., 2008). During the pilot test, students were asked to select what they thought was the correct response to the item and to answer follow-up questions about the item. These follow-up questions provided information about how well the item was performing for the target audience and whether or not the students were using the intended knowledge to answer the items. Students were asked to explain why they chose or rejected each answer choice; describe anything they found confusing; identify words with which they were unfamiliar; and comment on the helpfulness of diagrams, pictures, and tables. Items were then revised or eliminated based on the students' and panelists' feedback so that the final set of items could be considered a valid and reliable measure of the students' understanding of the energy concept.

Two Levels of Tests Used

Two levels of tests were constructed—basic-intermediate level tests and intermediate-advanced level tests. The intermediate-advanced level tests included all of the items (basic, intermediate, and advanced) and the basic-intermediate tests included items from the basic and intermediate levels only. Advanced level items could not be used with elementary school students because those items covered ideas that were too difficult for them and used terminology that would be unfamiliar to them. Matrix sampling, using multiple test forms at each level, was used so that

Table 4
Demographic information for the participants

	Elementary	Middle	High	Total
Grades	4–5	6–8	9–12	4–12
Number of students	2967 (14%)	10207 (50%)	7377 (36%)	20551
Gender				
Male	48%	49%	46%	48%
Female	50%	48%	55%	50%
Race/ethnicity				
White	38%	48%	44%	45%
Asian	7%	4%	7%	5%
Black	17%	11%	10%	11%
Hispanic	17%	19%	22%	20%
Two or more races/ethnicities	10%	10%	11%	11%
Primary language				
English	87%	88%	85%	87%
Other	11%	9%	13%	11%

we could test the wide range of ideas and levels of conceptual complexity that make up our hypothesized model. The basic-intermediate level tests included either 23 or 24 items, and the intermediate-advanced level tests included either 31 or 32 items. Linking items that appeared on all forms were used so that item characteristics could be compared across forms. Each of the 359 items was answered by an average of 1,605 students. Students in grades 4 and 5 took the basic-intermediate tests, and students in grades 6 through 12 took the intermediate-advanced tests. Students were given one class period to complete the test. For each item, students were asked to choose one answer; students who chose more than one answer did not receive credit for that item. Items were scored dichotomously.

Participants

Teachers from across the U.S. were recruited by email to participate in the study. All of the teachers who registered and had obtained administrative approval in their school district were sent testing materials; 328 teachers (about 82%) administered the tests to their students. A total of 21,061 students were tested in the study, but only data from the 20,870 students who responded to six or more items were analyzed. Students with highly unexpected responses were excluded as described below in the Rasch Analysis section. This resulted in a final sample of 20,551 students. Table 4 provides demographic

Table 5
Summary of Rasch fit statistics

	Item			Student		
	Min	Max	Median	Min	Max	Median
Standard error	0.02	0.11	0.06	0.37	1.93	0.40
Infit mean-square	0.84	1.27	0.99	0.44	2.16	0.99
Outfit mean-square	0.72	1.33	0.99	0.23	5.16	0.97
Point-measure correlation	0.00	0.53	0.34	−0.93	0.91	0.32
Separation index (reliability)		11.69 (0.99)			1.40 (0.66)	

information by grade level. Across the entire sample, 48% of the students were male and 50% were female; about 45% of the students were white, 11% were African American, 5% were Asian, 20% were Hispanic, and 11% reported two or more races/ethnicities; about 11% of the students stated that English was not their primary language. The sample included students from schools in 42 U.S. states and Puerto Rico. Elementary students (grades 4 and 5) made up 14% of the sample, middle school students (grades 6 through 8) 50%, and high school students (grades 9 through 12) 36%. All of the students were studying science but not necessarily physical science at the time of testing.

Rasch Analysis

WINSTEPS (Linacre, 2016) was used to estimate Rasch student and item measures. The data's fit to the Rasch model was evaluated using the separation indices, infit and outfit mean-squares, standard errors, and point-measure correlations. Separation indices of greater than two were considered acceptable (Wright & Stone, 2004), as were infit and outfit mean-square values between 0.7 and 1.3 (Bond & Fox, 2007). All point-measure correlations had positive values.

Initial analysis of the fit statistics showed that there were 10 items with outfit mean-square values outside of the acceptable range of 0.7–1.3 (Bond & Fox, 2007). The outfit statistic was used because it is unweighted and, therefore, sensitive to outliers. An investigation of the student response patterns for these items was conducted starting with the item with the highest outfit mean-square value. Data from 510 students with highly unexpected responses were removed from the data set, resulting in a total of 20,551 students in the final sample and a final set of items with infit and outfit statistics that were all within the acceptable range. Table 5 summarizes the fit statistics for both the items and the students. The reliability of the item measures was 0.99, and the item separation index was 11.69. The reliability of the student measures was 0.66, and the separation index was 1.40. This lower separation index and reliability for the student measures can be explained by the fact that students answered between six and 24 test items (basic-intermediate test) or between six and 32 test items (intermediate-advanced test) because of our use of matrix sampling. Therefore, there was less information available to estimate the student measures than was available to estimate the measures of item difficulty, where each item was answered by about 1,600 students. As a result, the student measures had a lower reliability and higher standard errors. Because our interest is in item difficulty and not individual student performance, the lower person reliability is not a concern for this study.

Item measures were then used as an indicator of where an idea fell on the learning progression (Wilson, 2009; Black, Wilson, & Yao, 2011). Easier ideas were assumed to come earlier in the progression, and more difficult ideas were assumed to come later in the progression. Wright maps (Wilson, 2005) were generated to visually represent where each item fell on the range of item difficulties. This information was then used in our analysis of each level of the progression, as described below. On a Wright map, students' performance level is shown on the left-hand side, and item difficulties are shown on the right-hand side. Easier items and less knowledgeable students are shown toward the bottom of the map, and harder items and more knowledgeable students are shown toward the top of the map. When a student's performance level equals an item's difficulty, the student has a 50% chance of responding to that item correctly.

Item Difficulty by Energy Category and Idea

To answer Research Questions 1 and 2 about the validity of the progression of the four energy categories and the 14 more specific energy ideas within those four categories, one-way analysis of variance (ANOVA) was performed to compare the mean item difficulties of the different energy

categories and ideas. Bonferroni post hoc tests were used to determine which energy categories and which energy ideas were more difficult than others.

Item Difficulty by Conceptual Complexity Level for Each Idea

To answer Research Question 3 about the validity of the three cognitive complexity categories, Kendall's tau correlation coefficients were calculated to assess the relationship between the difficulty of the items and the items' conceptual complexity level. When non-significant or negative correlations were found, an item-level analysis of the Wright map was conducted to determine if a different progression would better fit the data.

Student Performance by Grade Band

To answer Research Question 4, about how students are currently performing at each grade band with respect to the ideas in the learning progression, an analysis of covariance (ANCOVA) was performed, with demographic variables, including students' gender, race/ethnicity, and whether or not English was their primary language included as covariates. To control for differences in instructional focus across the country, the state that students came from was also included as a covariate. Bonferroni post hoc tests were used to determine if students in one grade band significantly outperformed students in another.

Results

Item Difficulty by Energy Category and Idea

To investigate the progression of item difficulty for the energy categories and ideas (Research Questions 1 and 2), we calculated the average Rasch difficulty of the items that were aligned to each idea (see Table 6). One-way ANOVA revealed statistically significant differences in the means of the 14 ideas, $F(13, 345) = 3.70$, $p < 0.001$. Bonferroni post hoc tests showed that the Chemical Energy items were significantly more difficult than the Elastic Potential Energy items, the Radiation items, and the Kinetic Energy items; and the

Table 6
Difficulty of energy ideas

			Rasch Difficulty				
Energy Ideas	Energy Category	# of Items	Min.	Max.	Mean	SD	
Elastic potential energy	Forms	18	−2.09	1.46	−0.45	0.87	Less
Transferring energy by radiation	Transfer	26	−1.45	1.33	−0.31	0.62	difficult
Kinetic energy	Forms	40	−1.49	2.08	−0.23	0.86	|
Thermal energy	Forms	40	−1.30	0.77	−0.17	0.52	|
Energy transformations	Forms	29	−0.79	0.72	−0.09	0.45	|
Gravitational potential energy	Forms	35	−1.26	1.21	−0.04	0.86	|
Transferring energy by forces	Transfer	23	−1.48	1.20	0.00	0.62	|
Dissipation & degradation	Diss.Deg.	21	−1.52	1.44	0.01	0.81	|
Transferring energy by sound	Transfer	12	−0.57	0.79	0.01	0.44	|
Transferring energy by conduction	Transfer	26	−1.19	1.83	0.08	0.72	|
Transferring energy by convection	Transfer	17	−0.52	2.24	0.26	0.73	|
Transferring energy electrically	Transfer	11	−0.95	1.12	0.34	0.64	↓
Chemical energy	Forms	28	−1.56	1.63	0.39	0.83	More
Conservation	Cons.	33	−1.01	1.98	0.50	0.82	difficult

Table 7
Mean item difficulty by energy category

Energy Category	# of Items	Mean Rasch Difficulty	SD
Energy forms and transformations	190	−0.09	0.73
Energy transfer	115	0.02	0.71
Energy dissipation and degradation	21	0.01	0.81
Energy conservation	33	0.50	0.82

Conservation items were significantly more difficult than the items aligned to Elastic Potential Energy, Radiation, Kinetic Energy, and Thermal Energy.

When the items were grouped into the four broad energy categories, we were not able to replicate at the $p < 0.05$ level of significance the finding of a progression from Energy Forms and Transformations, to Energy Transfer, to Energy Dissipation and Degradation, to Energy Conservation as suggested by previous research (see Table 7). One-way ANOVA and Bonferroni post hoc tests showed that only the Conservation idea was significantly more difficult than ideas about Energy Forms and Transformations and ideas about Energy Transfer ($F(3, 355) = 5.88$, $p < .01$). However, when all 14 energy ideas are rank ordered by average difficulty (see Table 6), it is clear that with only few exceptions, items testing students' understanding of the forms of energy tend to be easiest, items testing their understanding of energy transfer are more difficult, and items testing their understanding of conservation are the most difficult.

Item Difficulty by Conceptual Complexity Level for Each Idea

We also hypothesized that items testing the *basic* conceptual complexity level would be easier than the *intermediate* level items, and that the intermediate level items would be easier than the *advanced* level items. That is, we expected to see a positive correlation between item difficulty and conceptual complexity level. The Kendall's tau correlation coefficient for all items for all the ideas combined showed that the difficulties do significantly correlate with level, $\tau = 0.407$, $p < 0.001$. The mean difficulty in logits was −0.81 for basic level items, −0.05 for intermediate level items, and 0.45 for advanced level items (see Table 8). This means that our progression of conceptual complexity that begins by using concrete, easily observable energy phenomena and progresses to more complex and abstract ideas was confirmed.

To explore the validity of the progression of conceptual complexity within each idea (Research Question 3), correlations between item difficulty and level of conceptual complexity were calculated for 13 ideas (see Table 9). (Note that one of the original 14 ideas had only one level.) Kendall's tau correlation coefficients showed statistically significant correlations at the 0.05 level or better between item difficulty and conceptual complexity level for all but two of the 13 ideas. Only Thermal Energy and Transfer of Energy by Forces did not show significant correlations between the levels of cognitive complexity and item difficulty.

Table 8
Mean item difficulty by level of progression

Conceptual Complexity Level	# of Items	Mean Rasch Difficulty	SD
Basic	51	−0.81	0.62
Intermediate	190	−0.05	0.64
Advanced	118	0.45	0.63
	Kendall's τ	0.407	$p < 0.001$

Table 9
Mean item difficulty by level of progression for each idea

	Mean Rasch Difficulty			Correlation	
Energy Ideas	Basic	Intermediate	Advanced	Kendall's τ	p
Elastic potential energy	−1.51	−0.16	−0.12	0.523	<0.01
Transferring energy by radiation	−1.24	−0.51	0.05	0.579	<0.001
Kinetic energy	−0.99	−0.39	0.78	0.493	<0.001
Thermal energy	−0.99	−0.12	−0.07	0.167	n.s.
Transferring energy by forces	−0.52	−0.04	0.51	0.263	n.s.
Gravitational potential energy	−0.46	−0.01	0.49	0.320	<0.05
Dissipation & degradation	−1.07	0.40	0.50	0.527	<0.01
Transferring energy by sound	−0.51	−0.27	0.27	0.596	<0.05
Transferring energy by conduction	−0.57	0.06	0.80	0.418	<0.01
Transferring energy by convection	−0.06	−0.03	0.62	0.445	<0.01
Transferring energy electrically	−0.72		0.58	0.572	<0.05
Chemical energy	−1.32	0.61	0.79	0.461	<0.01
Conservation	−0.43	0.01	0.80	0.477	<0.001

Student Performance by Grade Band

To determine how well students performed at three different grade bands (Research Question 4), ANCOVA was used to perform a cross-sectional analysis of the students' performance by grade band, controlling for gender, race/ethnicity, whether or not English was their primary language, and the state where they went to school. Table 10 presents the F-ratios and degrees of freedom for grade band and each covariate.

The estimated marginal mean student performance was −0.54 logits for the elementary school students, −0.46 logits for the middle school students, and −0.16 logits for the high school students (see Table 11). Using the score-to-measure table generated by Winsteps, these measures equate to a raw score of 38% percent correct for elementary school students, 40% correct for middle school students, and 47% correct for high school students. A Bonferroni post hoc test showed that high school students performed significantly better than middle school students, and middle school students performed significantly better than elementary school students.

Discussion

The purpose of this study was to validate a proposed learning progression for the energy concept. We began with the four categories of energy ideas that are generally thought to make up the energy construct. We then elaborated them into a total of 14 more specific ideas and specified

Table 10
Results from the ANCOVA

Source	df	F	p
Grade band	2	405.63	<0.001
Gender	1	12.30	<0.001
Race/Ethnicity	1	247.67	<0.001
English as primary language	1	126.67	<0.001
State	1	168.15	<0.001
Error	19789		

Table 11
Estimated marginal student means by grade band

			95% Confidence Interval	
Grade Band	Mean Student Measure	Std. Error	Lower Bound	Upper Bound
Elementary	−0.54	0.014	−0.56	−0.51
Middle	−0.46	0.008	−0.47	−0.44
High	−0.16	0.009	−0.18	−0.14

three levels of conceptual complexity for 13 of these ideas. We then tested the progression using multiple-choice assessment items that targeted each level of each energy idea. In this section we discuss the extent to which the findings support the hypothesized learning progression, the contributions to the field, and the limitations of our study.

Research Questions 1 and 2: Validating the Progression of Energy Categories

As shown in Table 7, our data did not statistically support the currently established learning progression from Forms of Energy, to Energy Transfer, to Energy Dissipation, to Energy Conservation, although we did find that energy conservation is the most difficult energy concept. Research by Lee and Liu (2010) suggest that energy conservation is more difficult because it requires a higher level of integration of energy ideas than the forms, transfer, and dissipation ideas require. However, a rank ordering of all 14 ideas by average difficulty (see Table 6), does show that items testing students' understanding of the forms of energy tend to be easiest, items testing their understanding of energy transfer are more difficult, and items testing their understanding of conservation are the most difficult.

Past studies on energy have grouped the different forms of energy together as a single concept, and they have treated the different mechanisms of energy transfer as a single concept. Because researchers have tended not to test these forms and transfer mechanisms separately, there is little research on the relative difficulties of the different forms of energy or the different energy transfer mechanisms. Although energy is a unitary concept, and we want students to appreciate that, it is also true that energy is manifested in different ways and in different contexts. By not considering these manifestations or forms separately, it is impossible to know which of them are more or less accessible and comprehensible to students. By treating them individually, our study was able to show that chemical energy is a particularly difficult idea when compared to the other forms of energy (see Table 6). We also found that the idea of energy being transferred by radiation is relatively easy when compared to the other mechanisms of energy transfer (see Table 6).

These findings could also help explain why the progression from forms to transfer, to dissipation, and finally to conservation is more complicated than had been previously thought. Because certain forms of energy are more difficult to understand (chemical energy), and some transfer mechanisms are easier to understand (radiation), the specific forms and the specific transfer mechanisms that are assessed will affect the difficulty of each of those categories and where they would fall on a learning progression.

Additionally, the idea of dissipation and degradation proved to be easier than was predicted. This is most likely due to how the idea was defined. We began by thinking that dissipation and degradation could be placed on the same continuum, beginning with dissipation and moving toward degradation. The first two levels of our progression dealt with dissipation, starting with the basic idea that objects tend to get warmer when involved in energy transfers, and then moving to the intermediate level idea that energy is released to the surroundings during energy transfers and transformed into thermal energy. The advanced level then moved to degradation (i.e., unless

prevented from doing so, energy will become more uniformly distributed). Neumann et al. (2013) have argued that the dissipation ideas are similar in difficulty to transformation and transfer ideas, and that degradation should be treated separately in the progression. When we did that, and separated the items in our study into two groups (items targeting dissipation ideas and items targeting degradation ideas), we found that the dissipation items had an average difficulty of −0.15 logits, and the degradation items were considerably more difficult with an average Rasch difficulty of 0.50 logits. The dissipation items are in the same difficulty range as the forms of energy and energy transformation items, and the degradation items are in the same difficulty range as the conservation items. This supports the idea that energy degradation comes later in the progression of difficulty, but dissipation belongs at the same difficulty level and can be learned at about the same time as energy transformation and transfer.

Research Question 3: Validating the Progression Within Energy Ideas

Kendall's tau correlation coefficients supported the validity of a progression of conceptual complexity as defined in Table 2 for 11 of the 13 ideas. As expected, items that required a basic/phenomenological explanation were easiest; items that used energy terminology and energy concept-based explanations were more difficult; and items that required more advanced and abstract energy concept-based explanations were the hardest. For example, items that test the basic level of understanding for conduction (a warmer object will get cooler when in contact with a cooler object) are, on average, easier than items that test the intermediate level (energy is transferred from the warmer object to the cooler object). And the intermediate level items are, on average, easier than the items testing the advanced level (energy is transferred by random atomic collisions) (see Table 8). As noted earlier, there were two (of 13) ideas for which this hypothesized cognitive complexity ordering did not hold. For thermal energy, there was no significant difference between the intermediate and advanced levels, and for transferring energy by forces, the means followed the expected trends, but the correlation coefficient was not significant. An analysis of the

```
       Students | Items
 2           .  +
             .  |
             .  |                                      Thermal Energy
             .  |
             .  |T
            .#  |                                                                              | All
            .# T|                                                            What does         | things
 1         .##  +      Basic     |     Intermediate     |   Advanced    |    TE depend on?     | have TE
          .###  |                |                      |               |                      |
          .###  |S               |                      |               |                      | EG0623
         .####  |                |                      | EG0274 EG0323 |                      | EG0832
        .##### S|                |                      |               | EG0664               | EG0452
        .#####  |                |                      | EG0344        | EG0601               |
     .########  |                | EG0484 RG2071 EG0743 | EG0614 EG0842 | EG0292 EG0555 EG0574 |
 0  .#########  +M               | EG0514 EG0506        |               | EG0591               | EG0641
   .##########  |                | EG0852 EG0565        |               | EG0581               |
 .############  |                | EG0463 EG0493 EG0474 |               |                      |
 .############ M|                | EG0525 EG0652        |               |                      | EG0384
 .############  |  EG0284 RG1801 |                      |               |                      |
  .###########  |S EG0314 EG0334 | EG0536 EG0543        |               |                      | EG0632
    .#########  |  EG0304        |                      |               |                      |
-1   .########  +                |                      |               |                      |
      .####### S|  RG0563        |                      |               |                      |
       .######  |  EG0263 RG0582 |                      |               |                      |
          .###  |T               |                      |               |                      |
          .###  |                |                      |               |                      |
            .#  |
            .# T|
-2           .  +
```

Figure 2. Wright map of the revised thermal energy progression. Each "#" is 136 students and each "." is 1–135 students.

items from these ideas was conducted to investigate the source of the deviation and to see if the progressions could be revised to better match the data.

Thermal Energy. An analysis of the Wright map for the items aligned to the thermal energy idea was performed to determine why the items did not fit the proposed progression (see Figure 2). We found that the items targeting the idea that the thermal energy of an object depends on the temperature of the object were clustered at the bottom of the map (less difficult) as expected. However, the items that targeted the idea that the thermal energy of an object also depends on the mass of the object clustered toward the top of the map (more difficult) not in the middle of the distribution as expected, even above those items that were assessing ideas originally considered to be part of the advanced level (targeting atomic/molecular ideas about thermal energy). In other words, although it is relatively easy for students to think that thermal energy depends on the temperature of an object, it is difficult for them to think that thermal energy also depends on the mass of an object. The data support a progression that starts with the idea that thermal energy depends on the temperature of an object, followed by the idea that thermal energy depends on the speed and number of atoms or molecules that make up the object, and ending with the idea that thermal energy depends on the mass of the object.

Based on this finding, we revised the original progression shown in Table 2 by separating the intermediate level statement into two parts. Items that assessed the idea that thermal energy depends on temperature were moved to the basic level, and items that tested the idea that thermal energy depends on mass were moved to the advanced level. Items that were originally considered at the advanced level were moved to the intermediate level. A number of items did not fit into this revised progression. Seven items asked students to select the factors that thermal energy depends on. These items cut across multiple levels of the revised progression and cluster on the map between the intermediate and advanced levels (see Figure 2). These items are useful for testing the thermal energy idea as a whole, but because they align to more than one of the revised levels, they were not included in the correlation coefficient calculation for the revised progression. Additionally, the six items that targeted the idea that all things have thermal energy were spread across the difficulty spectrum. While these items are also useful for testing the thermal energy idea as a whole, they are not useful for this revised progression and, therefore, were not included in the correlation coefficient calculation.

The mean Rasch difficulties for the items aligned to the revised levels are shown in Table 12. Kendall's tau correlation coefficient was calculated using the revised progression, and a large and significant value was found (see Table 12). We postulate that having a solid understanding of atomic/molecular ideas related to thermal energy, especially the idea that thermal energy depends on the number of atoms or molecules that make up the object, may be helpful for making sense of the idea that thermal energy depends on mass. If students understand that thermal energy increases as the number of atoms/molecules increases and that mass is a measure of the amount of matter/

Table 12
Mean item difficulty for the revised levels of progression for thermal energy

Level	# of Items	Mean Rasch Difficulty
1) Thermal energy depends on temperature	8	−0.91
2) Thermal energy depends on the speed & number of atoms/ molecules	14	−0.19
3) Thermal energy depends on mass	5	0.32
Kendall's τ	0.734	$p < 0.001$

```
        Students | Items
2              .  +
               .  |          Transferring Energy by Forces
               .  |
               .  |     Basic       | Intermediate  |   Advanced
               .  |T                |               |
              .#  |                 |               |
              .# T|                 |               | RG1652
1            .##  +                 | RG0773        | NG0523 RG1212
            .###  |                 | RG1132        |
            .###  |S NG0504         |               | RG1012
           .####  |                 | RG0742        | RG0763
          .##### S|                 | RG1172        |
          .#####  |                 | RG0753 RG1222 |
       .########  |                 | RG1162        |
0     .#########  +M                |               |
     .##########  |                 | RG1602        |
   .############  |                 |               |
   .############ M|                 | RG1202        |
   .############  |  NG0512 RG0962  |               |
    .###########  |S RG0972         |               |
      .#########  |                 |               |
-1     .########  +  RG0982         |               |
        .####### S|  RG0952         | RG1102        |
         .######  |  RG0992 RG1592  |               |
            .###  |T                |               |
            .###  |                 |               |
              .#  |
              .# T|
-2             .  +
```

Figure 3. Wright map of the revised transferring energy by forces progression. Each "#" is 136 students and each "." is 1–135 students.

number of atoms/molecules that make up the object, then they can reason that thermal energy increases as the mass increases.

Transferring Energy by Forces. We also looked more closely at the Wright map for the location of items targeting ideas about transferring energy by forces to reevaluate our progression of understanding for this idea (see Figure 3). In our original hypothesized progression, at the basic level students are expected to know that energy can be transferred by "pushes and pulls," both contact and non-contact. At the intermediate level, students are expected to know that energy is transferred mechanically whenever a force is exerted on an object and the object changes position. At the advanced level, students are expected to know that there are changes in kinetic and potential energy that take place when objects change relative position as a result of a force being applied. When we looked at the Wright map we saw that the items did not fit the proposed progression. They did, however, with only minor exceptions, cluster into three distinct groups. At the lower end of the map were items that target the idea that energy can be transferred by contact forces. In the middle of the difficulty range were items that target the idea that energy can be transferred by non-contact forces. Finally, the items at the higher end of the range target the idea that a change in position or shape is necessary for energy to be transferred by the force.

When we looked at specific items, two items fell outside their revised level groupings. Item NG0504 in the basic group (contact forces) was much more difficult than the other items at that level. This item, unlike the others, targets a very popular *force* misconception that it is a force (not energy) that is transferred during a collision. Previous studies have shown that many students think that a force becomes part of a thrown or hit object (Fischbein, Stavy, & Ma-Naim, 1989; McCloskey, 1983; AAAS, n.d.). Because this item does not directly test the idea that energy can be transferred by contact forces, but rather tests a misconception about forces, we determined that the

Table 13
Mean item difficulty for the revised levels of progression for transfer of energy by forces

Level	# of Items	Mean Rasch Difficulty
1) Transferring energy by contact forces	7	−0.92
2) Transferring energy by noncontact forces	10	0.17
3) Forces must act over a distance in order to transfer energy	5	0.90
Kendall's τ	0.713	$p < 0.001$

item was not appropriate for this analysis and was removed. Item RG1102 in the intermediate group (noncontact forces) is an item that asks students whether a stronger or weaker magnet will transfer more energy to a metal ball. This item was somewhat easier than the other items involving noncontact forces. This may be because some students are simply associating the terms "stronger" and "more" rather than using their understanding of transferring energy by noncontact forces. However, the Rasch fit statistics for this item indicate that it was functioning properly in relation to the full set of items, so the item was retained in the analysis.

Table 13 shows the average Rasch difficulty for the items in the revised levels of the Transferring Energy by Forces idea, after removing the item discussed above. The statistically significant Kendall's tau correlation coefficient supports this revised progression: Students first gain an understanding that energy can be transferred by contact forces, then that energy can be transferred by noncontact forces, and finally that energy will only be transferred when the force acts over a distance.

Research Question 4: Current Student Performance

To see where the students who participated in this study stood in their understanding of energy relative to the expectations described in our learning progression, we looked at how the average Rasch measures for the three grade bands compared to the average difficulties for the basic, intermediate, and advanced level items. First, with respect to the basic level items, we found that the average difficulty of the basic level items across the grade bands was −0.81 logits, which is below the average student measure at each grade band, even for students in the elementary grade band, who had an average student measure of −0.54 logits. (When the item difficulty is less than the student measure, the students have greater than a 50% of responding correctly to the item.) This suggests that students at all three grade bands have a good understanding of the basic level energy ideas. But when we look at how well students did on the intermediate level items, we see that the intermediate level items have an average difficulty of −0.05 logits, which is greater than the average measure for students at all grade bands (−0.16 logits for high school students, −0.46 logits for middle school students, and −0.54 logits for elementary school students).

It is not surprising that elementary school students have not met the intermediate level expectations. What is surprising is that neither the middle nor high school students have met those intermediate level expectations, even at the 50% level. For the advanced level items, with an average difficulty of 0.45 logits, we see that their difficulty level is considerably greater than the average measure for the high school students (and, of course, elementary and middle school students as well). For the elementary level, the findings are encouraging because they indicate that elementary students have, overall, a grade-appropriate understanding of energy. However, the findings for the middle and high school students indicate that their performance falls below expectations and that more support for the teaching and learning of energy is needed at these grade bands.

Study Contributions

This study contributes to a fuller awareness of the complexity of the energy concept and how student understanding of it progresses in school by providing an empirically validated learning progression for energy and associated assessments to measure students' growth along the progression from grade 4 to grade 12. Our learning progression supports previous research that says that students at all grade levels, even fourth graders, can gain an understanding of basic aspects of energy. Because our learning progression provides a more complete picture of the energy concept and the relative difficulty of ideas that comprise it than those that have been previously described, it provides a more detailed road map to guide instruction and assessment. In particular, it gives additional insight into which specific energy ideas are more difficult to learn than others and guidance concerning how each energy idea develops over time.

Study Limitations

There are a number of limitations to this study. First, the study was cross-sectional in nature. We collected data from students in different grades at a particular point in time instead of collecting data from the same sample of students over several time points. Second, we did not have control over what the students who participated in our study had been taught. We assumed that all of the ideas that were tested were in fact taught, at least to some extent at the level appropriate for that grade, and that differences in emphasis on one topic over another were evened out across the sample. In other words, we assumed that across the sample the students had approximately the same opportunity to learn each idea so that what we were measuring was the extent to which the students were able to learn what was presented to them and not differences in curricular emphasis. This seems like a reasonable assumption, but it is not something we have hard evidence to support.

We also assumed that on average the items that were used to test each part of the progression were comparable in structure, complexity, and alignment to the targeted learning goals across topics. In other words, we assumed that there were no significant differences in how students were assessed across the learning progression. This is also reasonable to assume, given that the entire set of items was designed according to the same item writing specifications, paying attention to the use of real world contexts, age-appropriate language, etc. Ideally, the learning progression would be further validated using a study where the same students were assessed over a period of years and students were given the opportunity to learn the targeted ideas in a more controlled way.

Conclusions

Our analysis of Rasch item measures revealed a more complicated picture of students' progression of understanding energy than a simple sequence of discrete steps from forms to transfer to dissipation to conservation. For example, we found that certain forms of energy are more difficult, certain energy transfer mechanisms are easier, and dissipation is easier than degradation. We also found that most of the energy ideas progress from a basic/phenomenological understanding, to energy-concept explanations, to more abstract energy-concept explanations. Our analysis of Rasch student measures showed that although the elementary school students that were tested had reached a grade-appropriate level of understanding of energy, the middle and high school students had not yet mastered the expected intermediate level of understanding.

These findings support the idea that students' understanding of energy concepts develops together, in an interconnected way, not in a rigid sequence that starts with forms of energy and ends with conservation. This suggests that aspects of each energy idea can be introduced at each grade band with increasing sophistication. For example, elementary school teachers should be able to focus on phenomenological ideas related to a range of energy ideas, followed by energy-related

concepts to explain real-world phenomena at the middle school level, and finally more sophisticated and abstract energy-related ideas including atomic/molecular explanations in high school.

Given the wide application of these energy ideas, it is critical that students understand them and how to apply them in different contexts and that educators understand the difficulties that students may have. The results of this study can inform and improve science instruction on the topic of energy by providing information about how the energy ideas progress in difficulty. Because these items are designed to be carefully aligned with a progression of understanding for energy ideas consistent with *Next Generation Science Standards* but not specifically aligned to any single curriculum or instructional approach, researchers and developers of curriculum materials will be able to compare the effectiveness of various materials and approaches with more precision and objectivity.

The authors would like to thank William J. Boone for his advice on Rasch analysis. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A120138 to the American Association for the Advancement of Science. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

AAAS Project 2061. (n.d.) AAAS Science assessment website. Retrieved from http://assessment.aaas.org

American Association for the Advancement of Science. (1993). Benchmarks for science literacy. New York, NY: Oxford University Press.

American Association for the Advancement of Science. (2001). Atlas of science literacy (Vol. 1) Author.

American Association for the Advancement of Science. (2007). Atlas of science literacy (Vol. 2) Author.

Black, P., Wilson, M., & Yao, S. Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. Measurement: Interdisciplinary Research and Perspectives, 9(2–3), 71–123.

Bond, T. G., & Fox, C. M. (2007). Applying the Rasch model: Fundamental measurement in the human sciences (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). Rasch analysis in the human sciences. Netherlands: Springer.

Breslyn, W., McGinnis, J. R., McDonald, R. C., & Hestness, E. (2016). Developing a learning progression for sea level rise, a major impact of climate change. Journal of Research in Science Teaching, 53(10), 1471–1499.

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). Learning progressions in science: An evidence-based approach to reform. New York, NY: Columbia University, Teachers College, Center on Continuous Instructional Improvement.

DeBoer, G. E., Herrmann-Abell, C. F., & Gogos, A. (2007). Assessment linked to science learning goals: Probing student thinking during item development. Paper presented at the National Association for Research in Science Teaching Annual Conference, New Orleans, LA.

DeBoer, G. E., Herrmann-Abell, C. F., Gogos, A., Michiels, A., Regan, T., & Wilson, P. (2008). Assessment linked to science learning goals: Probing student thinking through assessment. In J. Coffey, R. Douglas, & C. Stearns (Eds.), Assessing student learning: Perspectives from research and practice (pp. 231–252). Arlington, VA: NSTA Press.

DeBoer, G. E., Herrmann-Abell, C. F., Wertheim, J. A., & Roseman, J. E. (2009). Assessment linked to middle school science learning goals: A report on field test results for four middle school science topics.

Paper presented at the National Association of Research in Science Teaching Annual Conference, Garden Grove, CA.

DeBoer, G. E., Lee, H. S., & Husic, F. (2008). Assessing integrated understanding of science. In Y. Kali, M. C. Linn, & J. E. Roseman (Eds.), Coherent science education: Implications for curriculum, instruction, and policy (pp. 153–182). New York, NY: Columbia University Teachers College Press.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). Making sense of secondary science: Research into children's ideas. New York, NY: Routledge.

Duit, R. (2014). Teaching and learning the physics energy concept. In R. F. Chen, A. Eisenkraft, D. Fortus, J. Krajcik, K. Neumann, J. Nordine, & A. Scheff (Eds.), Teaching and learning of energy in K-12 education (pp. 67–85). New York: Springer.

Duncan, R. G., & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. Journal of Research in Science Teaching, 46(6), 606–609.

Fischbein, E., Stavy, R., & Ma-Naim, H. (1989). The psychological structure of naive impetus conceptions. International Journal of Science Education, 11(3), 327–336.

Forbes, C. T., Zangori, L., & Schwarz, C. V. (2015). Empirical validation of integrated learning performances for hydrologic phenomena: 3rd-grade students' model-driven explanation-construction. Journal of Research in Science Teaching, 52(7), 895–921.

Goldring, H., & Osborne, J. (1994). Students' difficulties with energy and related concepts. Physics Education, 29(1), 26–32.

Hadenfeldt, J. C., Neumann, K., Bernholt, S., Liu, X., & Parchmann, I. (2016). Students' progression in understanding the matter concept. Journal of Research in Science Teaching, 53(5), 638–708.

Heritage, M. (2008). Learning progressions: Supporting instruction and formative assessment. Washington, D.C.: CCSSO.

Herrmann-Abell, C. F., & DeBoer, G. (2014). Developing and using distractor-driven multiple-choice assessments aligned to ideas about energy forms, transformation, transfer, and conservation. In R. F. Chen, A. Eisenkraft, D. Fortus, J. Krajcik, K. Neumann, J. Nordine, & A. Scheff (Eds.), Teaching and learning of energy in K-12 education (pp. 103–133). New York: Springer.

Hokayem, H., & Gotwals, A. W. (2016). Early elementary students' understanding of complex ecosystems: A learning progression approach. Journal of Research in Science Teaching, 53(10), 1524–1545.

Lacy, S., Tobin, R. G., Wiser, M., & Crissman, S. (2014). Looking through the energy lens: A proposed learning progression for energy in grades 3–5. In R. F. Chen, A. Eisenkraft, D. Fortus, J. Krajcik, K. Neumann, J. Nordine, & A. Scheff (Eds.), Teaching and learning of energy in K-12 education (pp. 241–265). New York: Springer.

Lee, H. S., & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. Science Education, 94(4), 665–688.

Linacre, J. M. (2016). WINSTEPS Rasch measurement computer program. Version 3.92.1. Beaverton, Oregon: Winsteps.com.

Linn, M. C. (2006). The knowledge integration perspective on learning and instruction. In R. K. Sawyer (Ed.), The Cambridge handbook of the learning sciences (pp. 243–264). New York, NY: Cambridge University Press.

Liu, X., & Boone, W. J. (2006). Introduction to Rasch measurement in science education. In X. Liu, & W. J. Boone (Eds.), Applications of Rasch measurement in science education (pp. 1–22). Maple Grove, MN: JAM Press.

Liu, X., & Collard, S. (2005). Using the Rasch model to validate stages of understanding the energy concept. Journal of Applied Measurement, 6(2), 224–241.

Liu, X., & McKeough, A. (2005). Developmental growth in students' concept of energy: Analysis of selected items from the TIMSS database. Journal of Research in Science Teaching, 42(5), 493–517.

McCloskey, M. (1983). Intuitive physics. Scientific American, 248(4), 122–130.

National Research Council. (2007). Taking science to school: Learning and teaching science in grades K-8. Washington, D.C.: The National Academies Press.

National Research Council. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, D.C.: The National Academies Press.

Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. Journal of Research in Science Teaching, 50(2), 162–188.

NGSS Lead States. (2013). Next generation science standards: For states, by states. Washington, D.C.: The National Academies Press.

Osborne, J. F., Henderson, J. B., MacPherson, A., Szu, E., Wild, A., & Yao, S. Y. (2016). The development and validation of a learning progression for argumentation in science. Journal of Research in Science Teaching, 53(6), 821–846.

Rasch, G. (1980). Probabilistic models for some intelligence and attainment tests (Expanded ed.). Chicago: University of Chicago Press (Original work published 1960).

Roseman, J. E., Stern, L., & Koppal, M. (2010). A method for analyzing the coherence of high school biology textbooks. Journal of Research in Science Teaching, 47(1), 47–70.

Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. Journal of Research in Science Teaching, 35(3), 265–296.

Stern, L., & Ahlgren, A. (2002). Analysis of students' assessments in middle school curriculum materials: Aiming precisely at benchmarks and standards. Journal of Research in Science Teaching, 39(9), 889–910.

Todd, A., & Kenyon, L. (2015). Empirical refinements of a molecular genetics learning progression: The molecular constructs. Journal of Research in Science Teaching, 53(9), 1385–1418.

Wilson, M. (2005). Constructing measures: An item response modeling approach. Mahwah, NJ: Erlbaum.

Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. Journal of Research in Science Teaching, 46(6), 716–730.

Wright, B. D., & Stone, M. H. (2004). Making measures. Chicago: Phaneron Press.

Supporting Information

Additional Supporting Information may be found in the online version of this article.